

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-306697

(43) 公開日 平成7年(1995)11月21日

(51) IntCl.⁵

G10L 9/00

識別記号

M

庁内整理番号

FI

技術表示箇所

審査請求 未請求 請求項の数2 OL (全5頁)

(21) 出願番号 特願平6-98801

(22) 出願日 平成6年(1994)5月12日

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目1番6号

(72) 発明者 小西 宏志

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(72) 発明者 吉田 孝

東京都千代田区内幸町1丁目1番6号 日

本電信電話株式会社内

(74) 代理人 弁理士 草野 卓

(54) 【発明の名称】 欠落音声補間方法

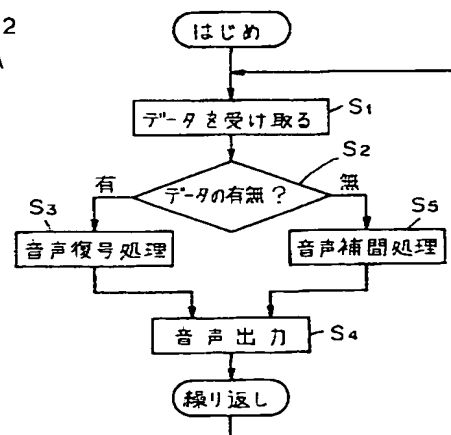
(57) 【要約】

【目的】 従来廃棄されていた遅延パケットの一部を有効に利用する。

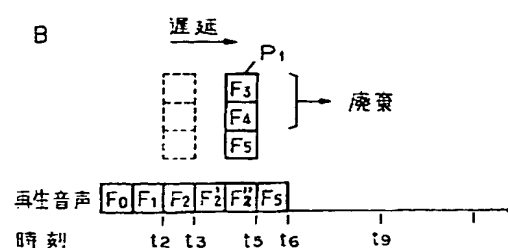
【構成】 図2Bに示すように、パケットP_iが再生開始予定時刻t₂に受信されていないと、従来と同様にその直前フレームF₂のデータを用いて音声信号F₂'、F₂''を時刻t₃、t₄に再生して補間するが、時刻t₄にパケットP_iが受信されると、本来それまでに再生されるべきフレームF₃、F₄を廃棄し、残りのフレームF₅を復号再生してそれまでの補間音声信号F₂'、F₂''から切り換えて再生音声信号F₅を出力する。

図2

A



B



【特許請求の範囲】

【請求項 1】 符号化された音声信号を複数フレームごとにまとめたパケットを受信し、

その受信したパケットから音声信号を復号、再生し、パケット欠落があると、その直前のパケットから音声信号を復号、再生して音声信号を補間する音声補間方法において、

各パケットの再生開始予定時刻から再生終了予定時刻までの間に到着した遅延パケットを検出する過程と、

上記検出した遅延パケットから、その遅延フレーム数だけ、そのパケット先端から除去した残りのフレームを抽出する過程と、

上記抽出した残りフレームから音声信号を復号、再生して、上記補間音声信号と切り換えて出力する過程と、を有することを特徴とする欠落音声補間方法。

【請求項 2】 上記音声信号の復号、再生を帰還形の音声合成フィルタを用いて行い、上記補間は、直前のフレームの復号フィルタ係数を上記音声合成フィルタに設定し、かつその音声合成フィルタを自己駆動として行うことを特徴とする請求項 1 記載の欠落音声補間方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は符号化された音声信号を複数フレームごとにまとめたパケットを受信し、その受信パケットから音声信号を復号、再生すると共に、パケット欠落時に、その直前のパケットから音声信号を復号、再生して音声信号を補間する音声補間方法に関する。

【0002】

【従来の技術】 音声信号をフレームごとに符号化し、その符号化音声信号を複数のフレームをまとめてパケットとして伝送する通信網においては、網内で生じる不規則な遅延、例えばパケットごとに異なるルートを通るために生じる異なる遅延のために、受信側にパケットが一定時間内に到達せずにパケットの欠落を生じることがある。そのため受信側において欠落パケット部分の音声信号を再生できなくなり、聴感上の音声品質の劣化が生じる。このパケット欠落問題を解決するために、従来よりいくつか欠落音声の補間方法が提案されている。その中でも、特許公開出願番号：特開平 5-88697 で提案されている適応予測符号化（APC）方式、マルチパルス駆動型線形予測符号化（MPC）方式もしくはコード駆動型線形予測符号化（CELP）方式等のピッチ分析を持つ音声符号化方式に適用可能な欠落音声補間方法を説明する。

【0003】 即ち APC 方式についてみると、これは送信側で、入力音声信号の短時間スペクトル包絡を表す短時間相関（フォルマント相関）を短時間予測器（フォルマント予測器）を用いて取り除き、更に有声音などにおいてピッチ周期ごとに波形が繰り返されるという長時間

相関（ピッチ相関）を長時間予測器（ピッチ予測器）を用いて取り除き、その結果の残差信号を量子化、符号化し、前記予測情報と共に送信する。

【0004】 受信側においては図 3 に示すように、入力端子 11 からの入力符号化音声信号は多重分離回路 12 で残差信号量子化情報と、長時間予測情報と、短時間予測情報とに分離され、これら分離された残差信号量子化情報から残差信号復号器 13 で残差信号が復号され、長時間予測情報から長時間予測係数復号器 14 で長時間予測係数が復号され、短時間予測情報から短時間予測係数復号器 15 で短時間予測係数が復号される。復号された残差信号はスイッチ 16 を通じて加算器 17 及び長時間予測器 18 よりなる長時間合成フィルタ 19 へ供給され、長時間合成フィルタ 19 の出力は加算器 21 及び短時間予測器 22 からなる短時間合成フィルタ 23 へ供給される。復号された長時間予測係数は長時間予測器 18 に設定され、つまり長時間合成フィルタ 19 のフィルタ係数が設定され、また復号された短時間予測係数は短時間予測器 22 に設定され、つまり短時間合成フィルタ 23 のフィルタ係数が設定される。

【0005】 従って復号された残差信号に、除去された長時間相関が長時間合成フィルタ 19 で付加され、更に除去された短時間相関が短時間合成フィルタ 23 で付加されて、音声信号が再生されて出力端子 24 へ出力される。入力符号音声信号が所定時間内に到達しないと、つまりパケットが欠落すると、端子 25 からの欠落情報によりスイッチ 16 がオフとされ、長時間予測器 18、短時間予測器 22 には直前のパケットにおける長時間予測係数、短時間予測係数がそれぞれ設定される。従って直前のパケットの長時間合成フィルタ 19 の出力が帰還作用により自己駆動として用いられて合成出力され、これが短時間合成フィルタ 23 で短時間合成され、欠落パケットに対する音声信号の補間が行われる。

【0006】 1 フレームの時間は符号化方式により異なるが通常 5～30ms であり、パケット長も各種あるが、例えば図 4 A に示すように 3 フレームごとに 1 パケットとして送出され、従って例えばフレーム $F_3 \sim F_5$ の符号化音声よりなるパケット P_i が、その直前のパケット P_0 の最後のフレーム F_2 の復号、再生が終了する時点 t_3 より前に受信されると、フレーム F_2 の復号、再生音声信号にパケット P_i の復号、再生音声信号を連続させることができる。しかしパケット P_2 のように、パケット P_i の復号、再生が終了する時刻 t_6 よりも前であるパケット P_2 の再生予定時刻、例えば t_5 までに受信されないと、パケット欠落として、その直前のフレーム F_2 を再生してその音声信号 F_2' , F_2'' , F_2''' を代用して補間処理を行っている。

【0007】 つまり、従来における処理手順は図 4 B に示すように、再生予定時刻までに音声パケットが受信されるかを監視し（S1）、受信されると、そのパケット

3

をバッファリングし (S_2)、再生すべき時刻になるとバッファリングしたパケットを復号、再生処理し

(S_3)、音声信号を出力して次のパケット受信を待つ (S_4)。パケットが再生予定時刻までに受信されないと、直前のフレームを用いて再生処理して補間音声信号を作って出力する (S_5)。

【0008】

【発明が解決しようとする課題】先に述べたように、従来においては再生予定時刻までにパケットが受信されないと、パケット欠落として処理していた。このため図4 Aに示すように、パケット P_2 が時刻 t_6 の直後に受信された場合は、フレーム F_6 の復号、再生には間に合わないが、フレーム F_7 、 F_8 は復号、再生して音声信号として用いることができるに拘わらず廃棄されていた。つまり、従来においては本来は良品質の音声信号が得られるべき所を、過去の情報による低品質の音声信号で補間していたという問題があった。

【0009】

【課題を解決するための手段】この発明によれば、各パケットの再生開始予定時刻から再生終了予定時刻（正確にはそのパケットの最後のフレームの再生開始予定時刻）までの間に到着した遅延パケットを検出し、この検出した遅延パケットから、その遅延フレーム数だけ、そのパケット先端から除去した残りのフレームを抽出し、この抽出した残りフレームから音声信号を復号、再生して、これをそれまでの補間音声信号と切り換えて出力する。

【0010】

【実施例】この発明の実施例を図面を参照して説明する。この実施例では受信パケットの受信時刻に応じて処理を分ける。つまり、図1に示すように、音声パケットが受信されると (S_1)、その受信時刻 t を知り

(S_2)、この受信パケットの再生開始予定時刻 t_s と、再生終了予定時刻（正確にはそのパケットの最後のフレームの再生開始予定時刻） t_e とを参照する (S_3)。これら時刻 t_s 、 t_e は最初にパケットを受信した時に、その時刻と、パケット長（フレーム数）とから、そのパケットの再生音声に連続して次の再生音声信号を得るための次のパケットの再生開始すべき時刻を再生開始予定時刻 t_s として求め、またその再生終了すべき時刻を再生終了予定時刻 t_e として求め、以下各パケットごとに次々と再生開始予定時刻 t_s と再生終了予定時刻 t_e とを求める。

【0011】この再生開始予定時刻 t_s 及び再生終了予定時刻 t_e と受信時刻 t とを比較し (S_4)、 $t \leq t_s$

（再生開始予定時刻 t_s より前に受信）の場合は正常受信と判定して受信パケットをバッファリングし

(S_5)、 $t_s < t < t_e$ （再生開始予定時刻 t_s より後、再生終了予定時刻 t_e より前に受信）の場合は遅延パケットと判定し、 $t_e \leq t$ （再生終了予定時刻 t_e より

4

りに受信）の場合は利用不能パケットと判定し、このパケットを廃棄する (S_6)。バッファリングしたパケットはその再生開始予定時刻 t_s になると読み出して音声復号処理部へ供給される (S_7)。遅延パケットと判定された場合は未再生部、つまり再生処理に間に合うフレーム部分を抽出して音声復号処理部へ供給する

(S_8)。つまり、遅延パケットをフレームごとに分割し、その先端から現時刻（受信時刻 t ）のフレームまで、つまり本来はそれまでに再生されるべき部分を廃棄し、次のフレーム以後の残りフレームを未再生部として取り出して音声復号処理部へ供給する。音声復号処理部は例えば図3に示した構成の復号、再生処理を行う部分である。

【0012】この音声復号処理部では図2 Aに示すように、図1中におけるバッファリングしたデータ（パケット）を受け取るか、抽出した未再生部のデータを受け取ると処理をし (S_1)、データを受け取ることができれば、つまりデータが有れば、そのデータについて音声復号、再生処理を行い (S_3)、その音声信号を出力し (S_4)、データを受け取ることができなければ、つまりデータがなければそれまでの受信データを用いて音声信号を再生して音声補間処理を行い (S_5)、その音声信号を出力する (S_4)。

【0013】図3に示した手法の場合は、帰還形の音声合成フィルタにその直前のフレームの復号フィルタ係数を設定し、かつ音声合成フィルタを自己駆動とする。例えば図2 Bに示すように、パケット P_1 は再生開始予定時刻 t_s が t_2 であり、再生終了予定時刻 t_e が t_5 であるが、時刻 $t_s = t_2$ より遅れ、時刻 $t_e = t_5$ より前の時刻 t_4 に受信されたとする。時刻 t_2 にパケット P_1 が受信されていないため、その直前のフレーム F_2 の各再生音声 F_2' 、 F_2'' が時刻 t_3 、 t_4 に補間される。しかし、時刻 t_4 にパケット P_1 が受信されるため、遅延パケット P_1 からその先端より、現時刻 t_4 に再生されるべくフレーム F_4 まで、つまりフレーム F_3 、 F_4 を除去し、残りのフレーム F_5 が復号、再生処理され、音声信号 F_5 として時刻 t_5 に出力される。つまり、従来においては時刻 t_5 の音声信号は、前のフレームからの再生信号 F_2''' で補間されるが、この発明では正規の音声信号 F_5 が再生され、それだけ音声品質が向上する。

【0014】補間音声信号を得るには前記例に限らず、例えば特開平5-88697号公報中に示されている各種の手法を用いてもよい。

【0015】

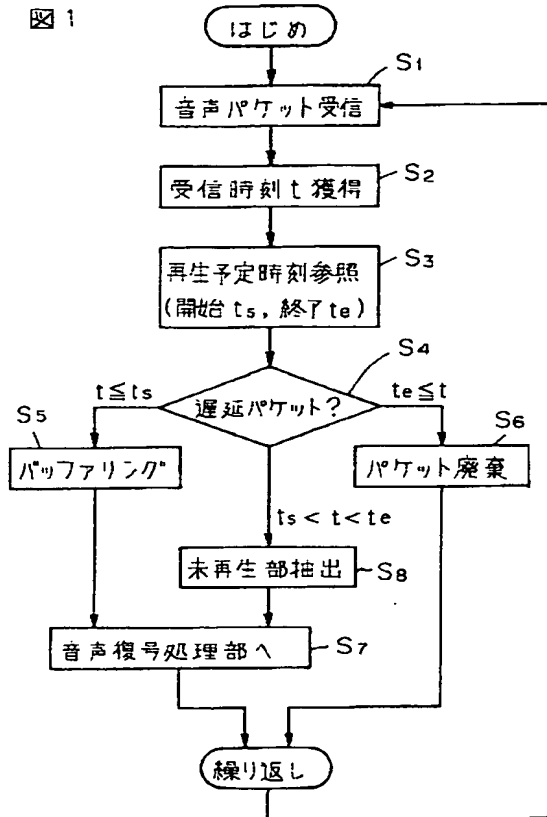
【発明の効果】以上述べたように、この発明によれば従来においては遅延パケット中の廃棄されていた利用可能部分を取り出して有効利用するため、その部分が補間音声信号よりも良品質となり、全体としても再生音声信号の品質が向上する。

【図面の簡単な説明】

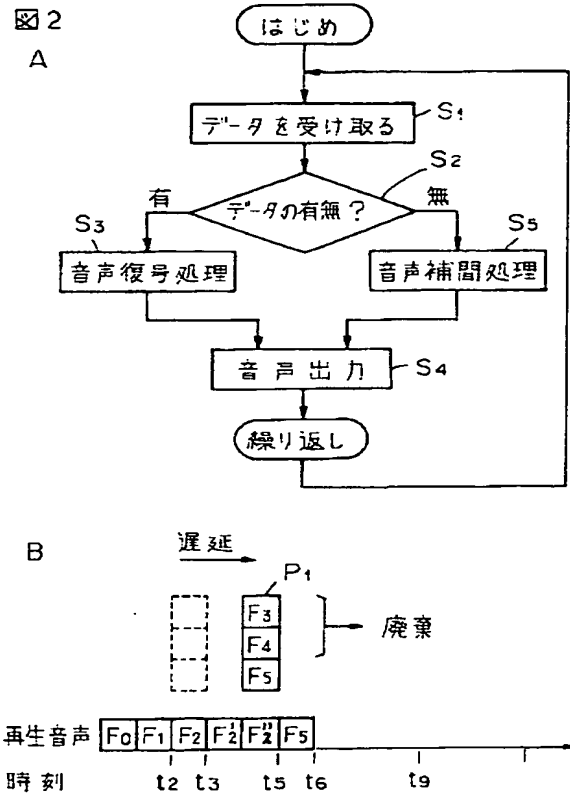
【図1】 この発明方法の実施例の要部における処理手順を示す流れ図。

【図2】 Aは受信データの音声復号処理と、音声補間処理との切り換え処理の手順を示す流れ図、Bはこの発明

【図1】

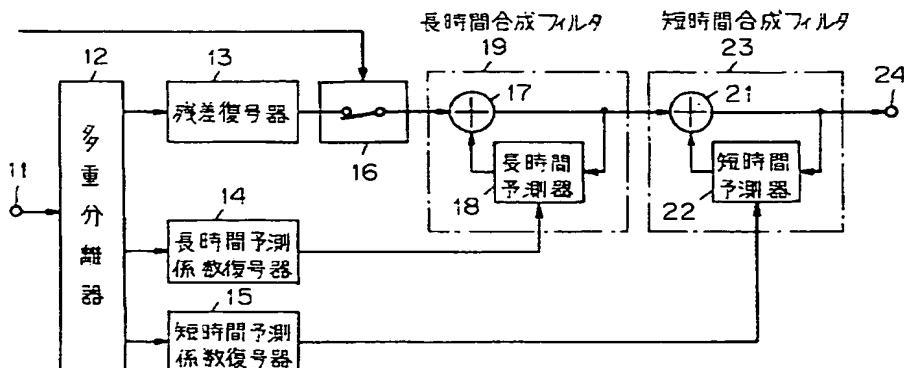


【図2】



【図3】

図3



【圖 4】

